AUTOMATED DETECTION OF PISTACHIO DEFECTS BY MACHINE VISION

T. C. Pearson, M. A. Doster, T. J. Michailides

ABSTRACT. *Pistachio* (Pistacia vera) *nuts with shell and kernel defects detract from consumer acceptance and, in some cases, may be more prone to insect damage, mold decay, and/or aflatoxin contamination. The objective of this study was to develop imaging algorithms to improve sorting of nuts with the following shell defects: oily stains, dark stains, adhering hull, and the following kernel defects: navel orangeworm (NOW) damage, fungal decay, and Aspergillus molds, all of which indicate risk of aflatoxin contamination. Imaging algorithms were developed to distinguish normal nuts from those nuts with oily stains, dark stains, and adhering hull as well as nuts having kernel defects. Image algorithm testing on a validation data set showed that nuts having oily stain, dark stain, or adhering hull could be distinguished from normal nuts with an accuracy of 98%. Removing nuts with oily stain, dark stain, and adhering hull will also remove 89.7% of nuts with kernel decay, 93.8% of nuts with Aspergillus molds present, and 98.7% of NOW positive nuts.*

Keywords. Aspergillus flavus, Aspergillus parasiticus, C*lassification, Image processing, Sorting.*

The United States is the world's second leading producer of pistachio nuts, producing approximately 160 million pounds per year at a wholesale value of nearly $2.00/lb. Nuts with certain shell discolorations such as those having oily stains, dark brown stains, and/or adhering hull detract from consumer acceptability and may indicate insect damage and/or aflatoxin contamination (Doster and Michailides, 1999; Pearson and Schatzki, 1998; Sommer et al., 1986). In particular, nuts with a narrow stain on the shell adjacent to their suture split indicate that it had a defect called an "early split" while growing on the tree. This suture stain is caused by a rupture of the hull along the suture, which can occur as early as six weeks before harvest. Normally, shells of pistachio nuts begin splitting open two to three weeks before harvest and the hull will stay intact, serving as primary protection for the kernel. However, on about 2 to 4% of the nuts, the hull will split open with the shell and form an early split. This opening leaves the kernel vulnerable to airborne mold spores and insect infestation. Early split nuts are more likely to be infested by insects and infected by the aflatoxin–producing molds, *Aspergillus flavus* and/or *A. parasiticus*, than normal nuts (Doster and Michailides, 1999; Sommer et al., 1986). It has been estimated that the incidence of aflatoxin contamination in early split nuts is approximately one in 500 and the contamination incidence among all harvested nuts is between one in 21,000 to 25,000 nuts (Sommer et al., 1986). In addition to suture stains, shells of early split nuts may have other dark stains, oily stains, and adhering hull (fig. 1). Many nuts with an oily appearing stain have insect damage, which might also contribute to aflatoxin contamination. Presumably, oil from the damaged kernel is absorbed into the shell, giving it the oily appearance (Doster and Michailides, 1999).

The pistachio industry currently uses electronic color sorting devices and hand sorting to remove pistachio nuts with shell discolorations. The color sorting devices inspect nuts at a rate of about 40/s and perform moderately well in removing heavily stained nuts but can erroneously reject good product due to the brown kernel exposed through the split shell. These devices are not effective for removing nuts with oily stains, suture stains or adhering hull, which frequently are associated with moldy kernels (Doster and Michailides, 1999). These defective nuts must be removed by hand sorting.

At present, a limited amount of research to develop automated sorting devices for pistachio nut defects has been reported. X–ray image histogram features and their spatial derivatives have been used for detection of insect–infested nuts (Casasent et al., 1998; Keagy et al., 1996a; Keagy et al., 1996b). These studies showed that, using X–ray imaging

Figure 1. Pistachio shell defect types from left to right: normal nut, adhering hull, oily stain, dark stain.

Article was submitted for review in January 2001; approved for publication by the Information & Electrical Technologies Division of ASAE in April 2001.

Reference to a company or product name does not imply approval or recommendation of that product by the U.S. Department of Agriculture or University of California to the exclusion of others that may be suitable..

The authors are **Tom C. Pearson,** ***ASAE Member Engineer,*** Agricultural Engineer, USDA–ARS, Grain Marketing and Production Research Center, Manhattan, Kansas, **Mark A. Doster,** Plant Pathologist**,** and **Themis J. Michailides,** Plant Pathologist**,** Department of Plant Pathology, University of California, Davis/Kearney Agricultural Center, Parlier, California. **Corresponding author:** T. C. Pearson, USDA–ARS, 1515 College Ave., Manhattan, KS 66502; phone: 785–776–2729; fax: 785–537–5550; e–mail: tpearson@gmprc.ksu.edu

18 to 43% of the insect–damaged nuts could be identified with only 1% false positives. In another study utilizing visible light, an image–based sorter was developed to inspect pistachio nuts for the characteristic "early split" stain adjacent to the shell suture (Pearson, 1996). This system utilizes three low cost linescan cameras placed around the perimeter of the nut and has a throughput of 40 nuts/s. Testing of this image sorter showed that it concentrates all of the aflatoxin from the main process stream into a reject stream comprising 2% of the sorted product (Pearson and Schatzki, 1998). However, the sorter was not found to be effective for removing pistachio nuts with dark stain, adhering hull, or oily stains, which often are associated with moldy kernels. Furthermore, the image sorter's ability to remove nuts having kernels with fungal decay, or *Aspergillus* molds, and insect damage has not been assessed. Thus, the objective of this study was to investigate the feasibility of adapting the image sorter (Pearson, 1996) for removing pistachio nuts with adhering hull, oily stains, dark stains, damage by navel orange worm (NOW, *Amyelois transitella*), fungal decay, or having *Aspergillus* present.

MATERIALS AND METHODS

LINESCAN IMAGE ACQUISITION

A California pistachio processor supplied samples of four pistachio nut categories with different shell defects (adhering hull, oily stains, dark stains, and normal nuts). Classification into each category was performed by quality control personnel at the processing plant. Normal nuts had no stains, or small stains comprising less than 10% of the shell surface area. From each of the four nut categories, 294 nuts were scanned in the image sorter (Pearson, 1996) for a total of 1176 samples from all four categories. Nuts were individually imaged by all three cameras at sorting line rates (approximately 1.5 m/s which corresponds to 40 nuts/s) as they passed through the image sorter. With nuts traveling at this rate, up to seven line scans from each camera could be obtained for each nut. The light sources used were halogen lamps (20W12V– 64425, Osram Sylvania, Danvers, Mass.) and the cameras were fitted with optical filters with a band pass band from 630 to 690 nm. These filters were found optimal for detecting brown stains on pistachio shells (Pearson, 1996). For this study, linescan images from all cameras were stored and processed off–line for algorithm development. Before capturing images from each defect category, 300 blank linescans from each camera were obtained. Using these data, the median pixel value at each pixel location was computed. This "median line scan" was then subtracted from all acquired images. As a result, linescan pixels representing the background had intensities near zero while pixels representing the nut ranged from 200 to 1000.

PROPERTIES OF NUTS

After imaging, nuts were taken to the Kearney Agricultural Center (Parlier, Calif.) where they were inspected for kernel decay, *Aspergillus* mold presence, and insect damage under a dissecting microscope (10x). Nuts were inspected for NOW developmental stage: filled pupa present, empty pupal case present, larval damage, or no evidence of NOW. Nuts were classified as having fungal decay if fungi were present and at least 1% of the kernel surface was discolored after the testa was removed. And finally, presence or no presence of *Aspergillus* sporulating structures on kernels was noted using a dissecting microscope (60x).

IMAGE FEATURE EXTRACTION

It was desired to develop discriminant functions (Huberty, 1994) to classify nuts into either a "reject class" or "accept class" based on a set of image features that could be extracted in real time. Several variations of which nuts comprised reject and accept classes are listed in table 1. Separate discriminant functions were computed for each of these reject and accept definitions. All discriminant functions perform two–way classifications. Oily stains, dark stains, and adhering hull were grouped into one reject class because they can all be present with normal nuts in processing streams and it is desired to eliminate all these defects with one sorting operation. Separate discriminant functions were developed to directly identify nuts with fungal decay, *Aspergillus* decay of the kernel, and insect damage.

Features used for the discriminant functions were limited to those that could be computed in real time with a sorting rate of 40 nuts/s. This constraint requires that features be limited to those that can be extracted from one linescan at a time because the process of storing a linescan in memory and re–calling it requires too much time with the hardware used (Pearson, 1996). Features were extracted from either the intensity linescan or the gradient linescan or both. The gradient image was computed by equation 1,

$$G_x = I_{(x-gap/2)} - I_{(x+gap/2)} \qquad (1)$$

where G_x is the gradient linescan value at location x, I_x is the intensity value of the pixel at location x, and gap is a distance in pixels.

Extracted features used in developing discriminant functions were:

1. The number of regions in all linescans from all cameras where all pixels exceeded an intensity and gradient threshold. Three different intensity thresholds were used (400, 500, or 600), three different gradient thresholds were used (10, 20, 30), and these gradients were computed with three different gaps (2, 6, 10 pixels). In order for a set of pixels to be considered a region, the length had to exceed either 5 or 10 consecutive pixels. Thus, a total of 54 features of this type were computed for all images.
2. The same as (1), but where all pixels fell below the thresholds listed above.
3. Image histogram bins of the intensity image where the bins were 20 intensity levels wide.
4. Image histogram bins of gradient images. Separate histograms were computed from gradient images computed with a gap of 2, 6, and 10 pixels. All bins were 1 intensity level wide.

Table 1 Various definitions of reject and accept classes used for discriminant functions.

Reject Class	Accept Class
Oily stains, dark stains, adhering hull	Normal nuts
Insect damaged kernel	Undamaged
Fungal decay of kernel	No decay
Aspergillus present on kernel	*Aspergillus* not present

5. Image statistics of only those pixels representing the nut. A threshold value of 200 was used to separate the nut from the background. Statistics computed were the mean and variance of the intensity image, gradient image (2–pixel gap), gradient image (6–pixel gap), and gradient image (10–pixel gap). Also, the total number of pixels having an intensity above 200 was recorded.
6. One dimensional frequency spectra magnitudes from each linescan captured. One–dimensional frequency spectra magnitudes were computed for all linescans from all cameras using a 256–point discrete Fourier transform with a Hanning window (Brigham, 1988). The Fourier Transform magnitude from all linescans from one image were summed and saved as potential features.
7. All possible ratios of points in the summed frequency spectrum magnitude.

Feature Selection

A total of 8882 features were computed for each nut. Most of these features did not provide meaningful information for distinguishing reject from accepts classes. Furthermore, only a small number of these features can be computed in real time with the hardware available. The first step in selecting useful features was to determine if the means of a given feature were significantly different for the reject and accepts classes using Fisher's least significant difference test. Only those features that were found to have significantly different means with 95% confidence were considered further. This reduced the number of features to about 1000. Next, stepwise discriminant analysis (sls = 0.01, sle = 0.01) (SAS Institute Inc., Cary, N.C.) was used to select features for a discriminant function. Stepwise discriminant analysis was performed to select a small number of useful features using the odd–numbered samples as the calibration set. The classification accuracy was verified using the even–numbered samples as a validation set. Features used in the final discriminant function were determined as follows. The first feature selected in the stepwise process using the calibration set was used in a univariate discriminant function and the error rate of the calibration set recorded. Then, the second feature selected in the stepwise procedure was added and the error rate of the calibration set recorded. The addition of features was repeated until the calibration set error rate reached a minimum. This entire procedure was repeated for each definition of reject and accepts class listed in table 1. Once a minimum classification error was reached in the calibration set, the discriminant function was applied to samples in the validation set and the resulting error rate noted.

Results and Discussion

Table 2 shows the percentage of nuts with fungal decay, *Aspergillus* present, NOW pupa present, or NOW damage. Nuts with oily stains have very high incidences of NOW infestation, fungal decay, and *Aspergillus* infection (table 2). A total of 38 nuts were found to have *Aspergillus* molds. Of these, 30 (78.9%) were *Aspergillus* section *Nigri*, two (5.3%) were *Eurotium* spp, and six (15.8%) were unidentified species of *Aspergillus* as the colonies were too small.

Table 2 Percentages of nuts from each shell type having kernel defects.

Shell Defect Type	Percentage of Nuts with Kernel Defects[a]			
	NOW	Pupa	Fungal Decay	*Aspergillus*
Oily stained	96.3 a	62.9 a	89.5 a	7.5 a
Adhering hull	7.8 b	0.7 b	30.6 b	2.7 b
Dark stained	5.1 b	0.7 b	25.2 b	1.4 bc
Normal	0.0 c	0.0 b	0.0 c	0.0 c

[a] Results in each column with a different letter are significantly different at the P = 0.05 level.

The classification results on the validation set using discriminant functions generated from the calibration set are presented in table 3 for the different reject and accept classes. A false positive is defined as a nut from the accept class being classified as a reject, while a false negative is defined as a reject being classified as an accept. Good classification accuracy is achieved when reject and accept classes are based on shell defects rather than kernel defects. When distinguishing nuts with oily stains, dark stains, and adhering hull from normal nuts, a validation set false positive rate of 1.4% is achieved with a false negative rate of 2.3%. The features used for each discriminant function are listed in table 4. The classification accuracy is generally poor when the reject class contains nuts with a specific kernel defect, i.e. fungal decay, *Aspergillus*, and NOW. Fortunately, shell discoloration can serve as a strong indication of these three kernel defects. When the reject class is comprised of nuts having adhering hull, oily stains, or dark stains and the accept class is comprised of normal nuts, the predicted reject class from the validation set contains 89.7% of nuts with kernel fungal decay, 93.8% of nuts with *Aspergillus* present, and 98.7% of NOW positive nuts. Thus, the use of this technology to remove nuts with shell defects will not only improve product appearance but it will also improve kernel quality. Further testing on a much larger independent set of nuts would be required to verify actual sorting accuracy. However, using half the data to calibrate discriminant functions and the other half to validate classification results corresponds well to actual sorting classification results with independent sets of nuts (Pearson, 2000).

Conclusion

This study shows the feasibility of adapting the image sorter described in (Pearson, 1996) to distinguish pistachio nuts having oily stains, dark stains, and adhering hull (nuts likely to have poor kernel quality) from normal nuts with a validation set false positive rate of 1.4% and a false negative rate of 2.3%. Nuts classified as having oily stains, dark stains, or adhering hull comprise 98.7% of the nuts with NOW infestation, 89.7% of nuts with kernel fungal decay, and

Table 3 Classification results from the validation set only.

Reject Class	Accept Class	False Positives (%)	False Negatives (%)
Oily stain, dark stain, adhering hull	Normal nuts	1.4	2.3
Insect damage	Undamaged	9.2	14.5
Fungi decay	No decay	17.0	26.0
Aspergillus present	*Aspergillus* not present	47.0	26.0

Table 4 Features used for each discriminant function.

Reject Class	Accept Class	Features
Oily stain, dark stain, adhering hull	Normal nuts	Four frequency spectra magnitude ratios, one intensity image histogram bin, one region feature[a]
Insect damaged	Undamaged	Two frequency spectra magnitude ratios, five frequency spectra magnitudes mean of the gradient image (2–pixel gap)
Fungal decay	No decay	Two frequency spectra magnitude ratios, three intensity image histogram bin, two frequency spectra magnitudes, one region feature[b]
Aspergillus present	*Aspergillus* not present	Two intensity image histogram bins

[a] Number of regions of 5 pixels or more where all pixels have an intensity greater than 600 and a minimum gradient of 20 when computed with a 10–pixel gap.

[b] Number of regions of 5 pixels or more where all pixels have an intensity greater than 600 and a minimum gradient of 20 when computed with a 2–pixel gap.

93.8% of nuts with *Aspergillus* present. The image acquisition and feature extraction can all be executed in real time, at a throughput rate of 40 nuts/s, using the current hardware configuration of the image sorter.

References

Brigham, O. E. 1988. *The Fast Fourier Transform and its Applications.* Englewood Cliffs, N.J.: Prentice Hall.

Casasent, D. A., M. A. Sipe, T. F. Schatzki, P. W. Keagy, and L. C. Lee. 1998. Neural net classification of x–ray pistachio nut data. *Lebensmittel–Wissenschaft und –Technologie* 31(2): 122–128.

Doster, M. A., and T. J. Michailides. 1999. Relationship between shell discoloration of pistachio nuts and incidence of fungal decay and insect infestation. *Plant Disease* 83(3): 259–264.

Huberty, C. J. 1994. *Applied Discriminant Analysis*. New York: John Wiley and Sons, Inc.

Keagy, P. M., B. Parvin, and T. F. Schatzki. 1996a. Machine recognition of navel worm damage in X–ray images of pistachio nuts. *Lebensmittel–Wissenschaft und –Technologie* 29(3): 140–145.

Keagy, P. M., T. F. Schatzki, L–C Le, D. Casasent, and D. Weber. 1996b. Expanded image database of pistachio X–ray images and classification by conventional methods. In *Optics in Agriculture, Forestry, and Biological Processing II*, eds. G. E. Meyer and J. A DeShazer. Proc. SPIE 2907: 196–204.

Pearson, T. C. 1996. Machine vision system for automated detection of stained pistachio nuts. *Lebensmittel–Wissenschaft und –Technologie* 29(3): 203–209.

_____. 2000. Automated sorting of pistachio nuts with closed shells. *Applied Engineering in Agriculture* 16(1): 91–94.

Pearson, T. C., and T. F. Schatzki. 1998. Machine vision system for automated detection of aflatoxin contaminated pistachios. *J. of Agricultural and Food Chemistry* 46(6): 2248–2252.

Sommer, N. F., J. R. Buchanan, and R. J. Fortlage. 1986. Relation of early splitting and tattering of pistachio nuts to aflatoxin in the orchard. *Phytopathology* 76(7): 692–694.